THE PROVENANCES OF YOUR SIMULATION DATA

Michael Donn[1], Drury Crawley[2], Jon Hand[3], and Andrew Marsh[4]
[1]Centre for Building Performance Research, Victoria University, New Zealand
[2]US Department of Energy, Washington DC
[3]ESRU Stratchclyde University, Strathclyde, Scotland
[4]AEC-Simulation, Autodesk, Inc., Waltham

ABSTRACT

This paper proposes a set of principles for an international database of building materials that would meet Quality Criteria for use in building performance simulation. The proposal draws inspiration from the International Glazing Data Base, but suggests that this inspiration goes as far as the quality assurance goal, not the practice. Rather than propose new means of storage of existing information, or new means of guaranteeing the quality of that data, it proposes instead that all data used in simulation should have an associated quality score based upon the quality of the tests used to derive the data; the quality of the testing laboratory; and the reliability of the error estimates. It includes examples of how this form of Meta-Data might be included into a range of different Building Performance Simulation Packages, and how a commercial building product search engine might deliver the quality score as well as the data.

INTRODUCTION – THE ISSUES

"what is the light reflectivity of that acoustic tile?"; "what are the heat storage properties and light reflectivity of that brick?"; "what are the bi-directional solar distribution function (bsdf) properties of that blind system?". These and many other questions are already commonplace amongst designers working within a BIM-based design environment. General values published in text books are often used as 'the best available'.

In a market where BIM is increasing pressures on designers to use complex simulation tools early in the design process (Eisenberg, 2002), and where the Building Product Model is making this feasible, and also where clients are seeking firm answers early on to investment decisions that do not open them to risk of litigation, these text book values are no longer sufficient. They are unlikely to 'stand up in court' as best practice.

Provenance is defined in the dictionary as: (1) Source OR origin. In the context of its use in this paper, the secondary definition is of greater explanatory power: (2) the history of ownership of a valued object or work of art or literature. In order to meet common definitions of best practice the provenance of a piece of simulation data on an acoustic tile needs to record not only the measured sound absorption coefficients for each part of the audible spectrum, but also what the error limits are on the coefficients are; what test was used to determine them; and the name and address of the laboratory that produced the data. In an ideal world this acoustic tile sound absorption data would also be associated with files listing thermal and visual properties and each of these would include their own provenance attributes.

This paper proposes implementation of a set of Provenance Meta-Data tags to be associated with all building product data that is published for use in building performance simulation. These Meta-Data tags make it possible to rate the trust level of simulations based upon the provenance of the data used. This trust level has many potential uses from establishing potential error limits for a simulation to meeting a Code-defined minimum trust standard for compliance.

PRECEDENTS

In the simulation world at present, the most trusted data source is the International Glazing DataBase (IGDB)[ii]. This is a database containing high quality data where the provenance is well-documented[iv]. The trust in the IGDB is maintained through a stringent process of peer review and of independent testing. It even has a process for evaluating the testing laboratories.

What the trust level of the IGDB suggests is that an International Building Materials DataBase (IBMDB) with at least equivalent quality assurance processes is needed for all the other building materials whose properties are key to. However, scaling the IGDB process up to cover all objects that ought to be in an IBMDB seems impractical.

It is difficult to define a rigorous test of 'practicality'. However, examination of the process of a new testing laboratory adding a single glass item to the IGDB illustrates the issues that arise when one considers developing a definitive single database of all (glazing) products. At present, the IGDB documents approximately 2000 different types of glass. In order to be certified to test glass for addition to the IGDB, the testing laboratory must be certified. They must complete and submit a series of tests of a set of glass

samples and demonstrate that their results match within tolerance limits those held by the certifying authority. Then when their test results for a new glass product are submitted, there is a short period when they can be challenged by the other testing laboratories representing in many cases the individual glass manufacturer's commercial competition. It is impractical for any one organisation to develop a similar process that could handle the diversity of thermal simulation data – let alone one that also stored all the lighting and acoustic data as well.

However, the clear goal of any IBMDB that might be developed would be to at least match the stringency of Quality Assurance of the IGDB – CERTIFIED LABS; CHALLENGES to results facilitated; WELL_DOCUMENTED TESTING methods. It must also be EXTENSIBLE – easy to add further data types and data items.

The means by which these goals would most likely be achieved is an online database of building products which contains amongst other things the provenances of each piece of data. Developed from this would be a simple Quality Algorithm which could rate the provenance. Not only would the provenances be published, but the academic credibility of the testing organisations would also be rated. This algorithm for rating the provenance is then the Quality Score of the provenance. For example, if a manufacturer publishes their own in-house data, then the provenance rating for their test would be low. If the provenance contains data from an independently ranked laboratory, then the provenance Quality Score will have a high rating. Aggregating these provenance ratings would result in a score indicating the level of trust one should place in the data input to the simulation model.

With a system of this type in place, it becomes possible for code authorities to reduce the gaming of simulation based code compliance by requiring the use of the provenance trust score in all compliance documentation.

WHO CAN WE TRUST?

A crucial part of the IGDB is the academic credibility of the testing processes. These are published in reputable international journals and are thus peer reviewed in as open a manner as academic traditions allow. Any IBMDB must rely on the same standard of open, peer-reviewed academic credibility. However, none of this process is new. What is new is that this provenance information remains associated with the data itself right through the simulation documentation process.

Despite its very tightly organised 'approved laboratory' approach to the production of test data, the International Glazing Database developers still have a challenge system whereby other testing laboratories and the market competition producers of

other glass products may challenge a test result prior to its publication. Maintaining and refereeing such a system for all building materials and products is apparently an impossibly large project. A single central database of all products would be a) impossibly large; and b) require an unimaginably large army of regulators. What is required is a participatory system involving as many people as possible in building and maintaining the quality of the data in the IBMDB.
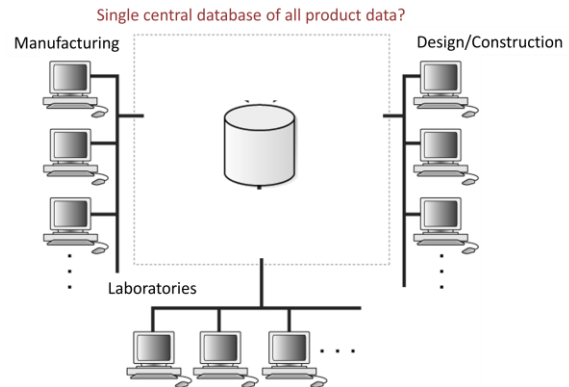

Figure 1 The Model of the IGDB - a single, central repository, carefully regulated certification of the laboratories and the manufacturers so users can trust the data
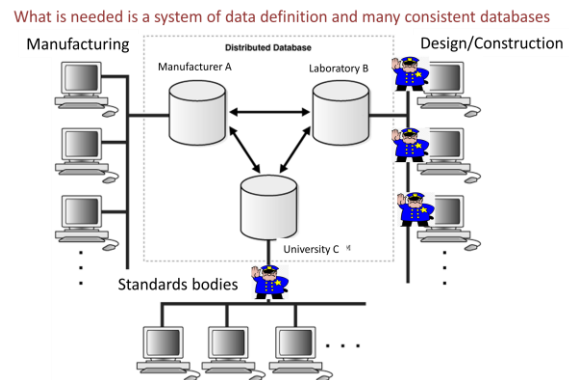

Figure 2 The model of the IBMDB: multiple databases; multiple systems of regulation: traditional academic refereeing; e-bay style user rating; IGDB style 'challenge'.

E-bay and other similar on-line auction houses have such a participatory system in their seller and buyer scoring. It is maintained by the community of users, with only a small team overviewing the entries. The overview is a policing of the potential for 'gaming' the ratings. An IBMDB scoring system would need to be maintained in much the same manner as the e-bay system – by the community of users. Then the number of potential referees – the users – has a chance of matching the number of products whose provenance needs to be assessed.

There are many competing laboratories testing building materials. There are many different testing methods. The policing of the cross-coupling between these is an impossible task for a single entity

following the IGDB model. It requires users to be demanding thermal properties with high quality scores matching the acoustic properties of their acoustic tiles. It requires those many consumers to be reporting publicly which products are well-documented and which not. It requires these consumers to be scoring the quality of the data.

In the same manner as E-bay and similar online auction systems rate the buyer and the seller, the IBMDB would also need to have a rating system for those who review the data. Everyone who has a product that was to be ranked is likely to be concerned that the 'Wikipedia' effect might become prevalent. They do not wish the rating of their product to be subject to some random student prank – or some politically motivated rating of its environmental impact – or some competitor's systematic attack. Allowing the product manufacturers to rate each and every person who provides a rating, and ensuring that only registered people with traceable addresses can submit ratings has proven to be a self-correcting model elsewhere. There is no reason to think it could not work in building product documentation.

In the future, this personal rating might well be linked to the academic publication process more formally. Thus, a person with a lengthy publication record in the same area as the test process, might be ranked higher than someone with a less relevant publication record.

POTENTIAL ERRORS DOCUMENTED

No measurement is exact. One of the critically important aspects of the provenance of a data item is that it is transformed from being a single number representing say R-value or reflectivity into a set of numbers representing, at a minimum, the 'typical' value and error bars associated with its measurement. Thus, the light reflectivity of a ceiling tile may well be reported as 0.8 – but the reality noted in the provenance is that it is most likely 0.8, but has a 95% chance of being between 0.75 and 0.85.

Armed with this data, the simulationist has the opportunity to produce another quality score for their simulation. At its most trivial level, this would involve running the simulation three times: once for all the data points at their lower bounds , once for all at their 'typical' values and once for all at their maxima. This would establish some crude estimate of the reliability limits of the simulation.

Some tools already include facilities to support parametric excursions across a range of data types and to identify the sensitivity of predictions to such bounds tests at a fine level of granularity [McDonald, 2002]. They are currently limited by relying on users opinions about the bounds to be tested.

A more sophisticated level of use would be tweak the simulation program to look for the most influential errors. Here the person checking the simulation, for design or code compliance would have a strong indication of what was the likelihood of the simulation being in error and by how much given likely variations in the input data. A probability analysis could also be produced of the likely range of variation in the resultant simulation prediction given likely variations in the input data.

IMPLEMENTATION – I: THEORY

There are some clear principles for the production and storage of provenance data. These are:

1) Data objects, not data points: in order to store the provenance in such a manner that it cannot be separated from its source, each piece of information must be stored and distributed in a format that ensures its associated provenance is stored as Meta Data in the same object;
2) To facilitate this process, the input data processor for a simulation tool must accommodate the provenance data;
3) The interface to the simulation program must make this association of the object's provenance with the actual input data for the simulation no more difficult than the current process of looking up a reflectivity or an R-value in a catalogue;
4) The Key Fields in a provenance for simulation data are:

<VALUE> The data value itself – reflectivity, R-value etc;

<UNITS> The units of the data value ('Fraction' for reflectivity; m^2 ^oK /W for R-value; if declared in this way, then machine- translatable;

<ERROR> The measurement imprecision – a data pair; if relevant for sampled data, the upper and lower bounds where 95% of the is observed.

<URI-M> location information for the manufacturer/producer of the object itself;

<URI-L> location information for the test laboratory;

<URI-L-Rate> location information for the certification from the relevant authority of the suitability of the test laboratory;

<URI-T> location information for the description of the test;

<URI-KEY> An identifier for the object that enables other data values to be stored / added.

The data fields in the above object are suited to storage in a database. Accessing the database would be easier if the object could be located digitally via the internet. As an XML file it could become a self-describing database. The method of delivery to the simulationist of the data describing the provenance of simulation data should be immediate, and ideally managed by the simulation interface.

In the short term, the method of delivery question could become mired in the debate about data formats

for the reliable storage and exchange of information about buildings. This paper is not the place for such a debate to be developed. Rather, it is an exploration of the potential implementation of this approach to managing the provenance of simulation data in the context of the following applications:

1) Autodesk® SEEK: an online building product information delivery system within which the scoring of data quality might be implemented;
2) ESP-r: a thermal simulation package which has well-developed links to CFD and Lighting analysis packages;
3) Autodesk® Ecotect™ is a building design and environmental analysis tool that has its own in-built early design algorithms and links to many industry standard full simulation products.

The next sections of this paper explore the potential application of this provenance approach within these three applications.
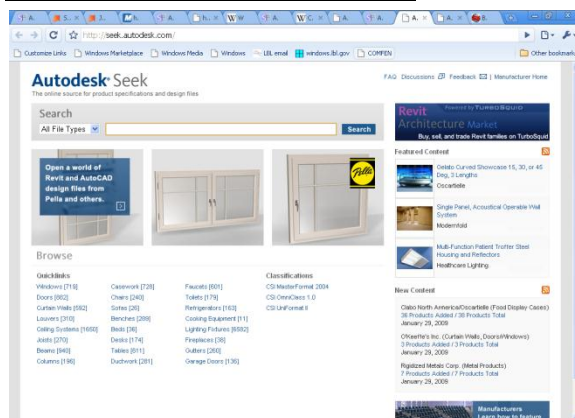
IMPLEMENTATION II – AUTODESK® SEEK – FINDING THE DATA

Figure 3 The Autodesk® Seek Web Site

In May 2008 Autodesk released a beta of Autodesk® Seek[viii], a web-based Architecture, Engineering and Construction (AEC) specific, 3D model and specifications search tool. Rather than yet another catalogue or model index the service is focused on exposing the model and specification catalogues of AEC suppliers. For architects and engineers the ability to quickly locate, access and reference specifications and 3D data could potentially reduce design development time and costs significantly.

The idea of an online product catalogue for AEC specifications is certainly not new (Yaman, 2000). However Seek is unique in that it is the first online product catalogue backed by a large company whose primary customer-base is not AEC suppliers. This independence establishes trust which is important because users do not want the relevancy of their search influenced by who is paying the bills, nor do they want a 'walled garden' where only products from selected (paying) suppliers are on show.

Consequently even though many supplier-backed catalogues exist, none can be considered the Google of the AEC world.

Seek has the potential of filling this 'Google' void because Autodesk's primary income is from people who make material purchasing decisions (architects, engineers and contractors, etc.) and not the suppliers themselves. This difference places Seek in the position of being able to design a catalogue that acts in the best interests of the search consumer. At the same time suppliers who do not take part risk "missing the boat" to given Autodesk's vast global audience.

The function of a web service
Seek advertises itself as a 'web service'. In simple terms, this is "a software system designed to support interoperable machine-to-machine interaction over a network" (W3 Corp). Seek's exposed functionality classifies, filters and links relevant files found during a search of the internet for building product data. Seek applies multiple classification systems to the data stored within its index. Typically, people looking for AEC content have specific contexts in mind and classification systems help define their respective boundaries. By limiting search results to a specific subset of the building industry the potential for finding what you are looking for increases dramatically.

According to the online talk by Mike Haley of Autodesk (Haley, 2008), they are developing on-the-fly algorithms that classify incoming data. They are allowing for 'foreign' semantic systems whether they be foreign languages or other classification systems such as the CBI classification system (CBI) used in New Zealand. Haley discusses the way these algorithms are intended to deal with semantic concepts – what people are really looking for when they do an AEC search. The search for example is being built to cope with and learn about hyponyms and hypernyms. A hyponym is a loose word that describes or categorises a more specific term – so a search for wood windows would find both oak and elm frame windows. Similarly Aluminium and Steel are both hypernyms of metal so a search for Aluminium windows would find all windows specifically listed as Aluminium and all windows just listed as metal, with no material specification.

This is the beginning of the means to search for AEC objects that match selected specifications as has been suggested in the past is a necessary means of delivery of building simulation data of high quality provenance. Haley describes the planned future functionality for users as drawing a square in a CAD modelling program and clicking the CAD program's search button for a matching window. His example is a casement window matching these drawn dimensions. It is not a great leap to add a required mix of window thermal conductance and visual

transmittance to the search terms. Because Seek is a search engine, it looks at data in online catalogues and delivers that information from whatever source.

Seek delivers whatever files it finds in the online catalogues it indexes. At present, these files include CAD files (including SketchUp and Microstation files, not just Autodesk files); pdf brochures; and thumbnail images. Again, in a world where the provenance of simulation data is significant, this system could deliver snippets of simulation files that accurately describe the material properties in a format suitable for inclusion in a simulation program. A simple implementation of the concept that could be begun immediately would be to deliver all the IGDB data in Window 5/6 database files that exist on the web in window manufacturers' catalogues.

Filtering

Seek offers a wide range of attributes to filter on once a category or basic search term has been defined. This mechanism enables quick culling of large sets of results to identify a couple of the most relevant models or specifications. In this regard Seek behaves more like an e-commerce site rather than a search engine because the emphasis is not on providing you 50 relevant suggestions but one or two specific answers.
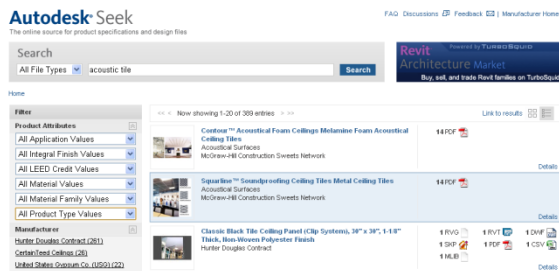

Figure 4 Seek's search results with its filtering system on the left

Seek automatically derives the filter attributes from indexed data, within a 'canonical taxonomy' being developed by Autodesk. An interesting aspect of the Seek auto-search function is the work being undertaken to enable the Seek system to "understand" what are reasonable ranges of properties for materials. This has great potential for enhancing the provenance of data. That the manufacturer's product data falls within a range of 'normal' values for such products is further reassurance that this simulation data is of reasonable quality.

When it comes to creating an accurate and timely index, blog search engines have demonstrated that the ability to push structured data to the search engine is far more efficient than using a conventional Web crawler approach. With this capability the very nature of the catalogue would shift from that of an online book to a living entity. If suppliers were able

to push availability details and news about a particular product into the index it would mean that any consumer of Seek data would also be able to utilise this information. For example:

An architect assigns a product specification from Seek to the Visual Transmittance of the windows on an office facade in their AutoCAD model. The Ecotect analysis has demonstrated that this is the critical value for the success of their daylight scheme. On making this assignment they select to be notified of important information on this product until the project is complete. This configures AutoCAD as a subscriber to the product-specific RSS feed on Seek. As any new information is announced by the supplier, for example it will be discontinued in December or a national safety test found it did not perform well under certain situations, then anyone opening the model would be alerted to this news. With its use of Atom feeds this is a potential future direction of the Seek product.

Exposing the Data

Crucial to the success of Seek is its web service component, i.e. the ability for other applications on the Web or desktop to use the data this service returns. Whilst Autodesk currently describe Seek as a 'web service' this is not the case in the contemporary sense (w3 Corp). Seek's value will increase exponentially once it makes the leap from a visual catalogue to a service which forms the functional backbone of desktop and web-based applications. The following two scenarios explore how a service-centric Seek could behave:

An architect working on an ArchiCAD model is about to make a design decision regarding a particular wall cladding. The ArchiCAD Seek plug-in recognises that this is the case because the user has selected the appropriate modeling tool and layer set. The plug-in queries Seek and returns a list of appropriate 3D models and their associated ESP-r thermal and daylighting data based on the properties of the project (a residential dwelling in a hot climate). The plug-in filters and orders this data to suit the architect's personal preferences - in this case supplies that are also from a sustainable forest. Without a single extra mouse click, Seek in partnership with the desktop software is able to present a reasonably intelligent set of currently available cladding options. This task, which could have taken hours of searching through conventional product catalogues and manual 3D modelling is completed in seconds.

A design team working on a medium sized office project in Sydney is having a discussion within their Project Intranet on appropriate light shelves to use for daylighting. The ESP-r analysis has shown that this design will provide high quality lighting at low energy use. None of the available products seem to suit the windy climate and ideal daylight reflectivity

requirements specified in the ESP-r analysis. The outstanding issue is recorded as something that needs attention later. The Intranet software constructs a Seek search query out of the issue's defined parameters and begins regularly checking Seek for potential matches. Weeks pass and the problem is forgotten about by the team. Then one afternoon the Intranet service issues an email informing the interested parties that a local supplier has just that morning started producing a new line of industrial strength fixtures which satisfy the design requirements.

In both these scenarios the use of web services transforms Seek from a user-initiated search tool to a context-aware information delivery service. The guarantee of a link to the provenance of the data describing the product provides the assurance that the product is real, and not some sales-hyped widget.

Provenance

General purpose search engines establish 'correctness' through the concept of *Google PageRank* (i.e. if it is linked to it is probably right). Unfortunately for the closed and competitive world of architectural design this concept cannot be applied even if it was possible for Autodesk to go crawling the design plans of AEC professionals to identify which models and specifications are referenced the most. However it would be feasible to deploy an opt-in system within Seek where users could identify models and specifications they made use of regularly. For example during the drafting of construction details the CAD program could notify Seek whenever specifications stored in the index were referenced by the designer. In practice this would be similar to Google's Web History as the aggregate, anonymised data returned would help assist others to identify popular, and therefore by logical extension trusted, models and specifications.

Beyond passive observation is the ability for users to directly feed into Seek's index their own opinions and content. For example much of the real value of the Amazon web experience is not the search results but the user reviews. Basic online specifications is one thing, but knowing that someone in a very similar situation as yours found the actual product did not measure up to expectations is considerably valuable.

Seek offers a potential provision of a level of trust and Quality Assurance not currently available within the building simulation world. It requires building into the service these user reviews and the provenance of each data item: the source of the measured data; the trust score for the organisation doing the measurement; links to the standards authority who defined the measurement method.

Leveraging the social

The "Web phenomenon" of the past three years has been the move towards social-centric networks (e.g. Facebook, MySpace and Twitter). AEC professionals subscribe to magazines and catalogues, visit interesting buildings and attend lectures because they want to know what their peers are up to. Seek could enable users to track what was 'in' and what was 'out'. Finding products that help a building type – say an office or a school - to achieve LEED accreditation in a particular climate or market would contribute significantly to quality assurance in simulation.

IMPLEMENTATION III – ESP-R: THERMAL SIMULATION

ESP-r has, for some time, included facilities to formally describe uncertainty within the data model e.g. the conductivity of a layer in a construction type or specific instances of that construction. It also has a number of statistical approaches to discovering the sensitivity of performance predictions to changes in a models thermophysical or geometric attributes and the ability to report on error bands and residuals (between different assessments) [ ]. ESP-r also includes facilities to manage databases via the interface rather than via ad-hoc editing of files so there is scope for these facilities to be extended to allow for external information to be incorporated into the infrastructure of a group.

The idea that the selection of the constituent parts of building includes clues as to a distribution of thermophysical values, and that our interpretation of performance should be guided by this rich set of attributes seemed like a good idea when conceived. Thus far it has been the case that having built the facility, almost no one has deployed it.

There are several possible causes for this and usual suspects would surely include:

- practitioners are less comfortable with statistics and concepts of uncertainty than tool developers

- practitioners are unsure how real products vary from default/published values

- data recovery and display techniques in most simulation tools do not do much justice to ideas that "there is more than one answer"

- facilities supporting database management in simulation tools is largely introspective and does not readily link to external sources

Each of these issues would benefit from the emergence of an IBMDB. An easy to access IBMDB would increase awareness of the range of entities and their attributes. It might also embolden the ESP-r development community to debate, influence and contribute to the form and composition of the IBMDB.

Given the limited resources for developing open source simulation tools, any step to evolve ESP-r to link to external sources carries risks. The first risk is of introducing dependencies that are difficult to maintain. External API which are complex will also delay implementation as will API which are based on proprietary standards.

The next issue for ESP-r is the interdependency of databases [ ]. For example, a façade construction may be linked to a materials database which includes solar and optical properties for a single layer as well as a separate database of optical properties for the whole construction. If there are conflicts between the entities (e.g. a different assumption about the thickness of the layer) it gets messy.   One consideration if the IBMDB were to be a single data store would be to design the data store so that potential conflicts are identified. A search engine fuelled IBMDB would require rules for identifying and establishing a hierarchy of reliability based upon the individual provenances of each conflicting value found.

Another issue for the design of the IBMDB data store is how to mix information derived from the IBMDB with existing ESP-r database entities, especially in the early stages when the data store is sparse. For elemental entities such as a colour or texture this is less an issue than for entities which have many attributes, some of which may not be complete.

Another observation is that it is difficult for some users to deal with verbosity of documentation. If there are a half dozen attributes of an entity and each attribute has a paragraph related to provenance and the user is confronted by several hundred entities there is a risk of users getting lost or losing interest. The IBMDB must find a means of presenting the overall product in a simple icon or thumbnail whilst allowing the user to drill down to the source of the reliability score.

Once simulation projects begin to be embedded with information from the IBMDB the data model of ESP-r needs to be extended to ensure that provenance is maintained within the model. And if a simulation group develops opinions about the IBMDB derived entities then it should be straightforward to record this within the model. Ideally, such additions should have an easy path back to the IBMDB data store so that others can benefit from new insights.

IMPLEMENTATION IV – AUTODESK® ECOTECT™: multi-criteria modelling

Much of the functionality required to implement the kind of provenance tracking, quality scoring and error margins described here can already be accommodated within Autodesk® Ecotect™. In v5.60, a flexible database was added in the file format to allow any amount of extra data to be added and associated with individual materials, equipment or zones. This, combined with some simple scripts and interactive wizards, allows Ecotect to access external data sources, store complex associative data within a model and then use it in calculations and custom generated reports.

Accessing External Data Sources

Ecotect script commands include functions such as `get.app.web.page` and `get.app.web.file` for accessing web data sources. Alternatively, LuaCOM can be used to invoke COM objects such proprietary libraries, remote databases or even tools such as Autodesk® Revit®, MS Access or MS Excel. Once the data is acquired, any fixed-format or XML content can be read using `get.app.web.param` or parsed directly into the model using LuaXML.

As shown in Figure 5, Ecotect allows users to create their own wizards with detailed user interfaces for accessing external sources of material or provenance data.
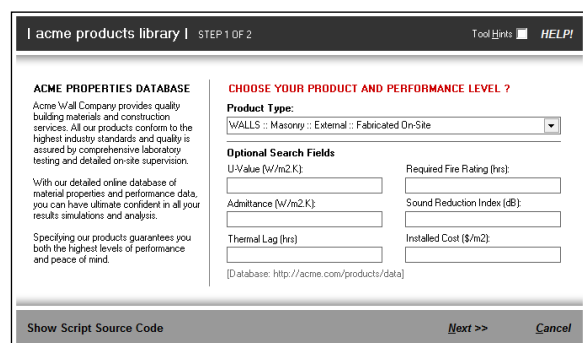

Figure 5 An example user-defined wizard providing a custom user interface to a remote material data source.

Storing Provenance Data

Additional material data can be stored in Ecotect as raw text or in the form of multiple token-value tables. The database itself is stored as embedded XML and can be manipulated manually using the interface provided in the PROJECT page, via script commands or even using an external XML editor.

To store provenance data, the user simply selects the materials they wish to import from the external data source and the script/wizard then creates new materials in the model and stores the additional provenance data - including all the Key Fields described above. The benefit of using a flexible database means that any number or type of extra URLs, indexes or references can be stored. This is important to facilitate systems in which different parts of a material's dataset are drawn from different sources. For example, the basic thermal properties of a masonry wall may come from a general properties data source whereas the colour, reflectance and internal emissivity may be drawn from a paint manufacturer's database.

Figure 6 A screenshot of provenance data associated with material 1031 in an Ecotect model.

Using Provenance Data

Once the information is in the embedded database, scripts or wizards can use it to either generate summary reports for all the materials in the model or to modify or replace data fields used in actual calculations. For example, a script could be used to perform a series of thermal performance calculations; iterating through provenance data to set material properties to their maximum and minimum error bounded values as well any number of interpolated values in between. At the end of each calculation, the script extracts the appropriate results and associates them with a record of the material properties used.
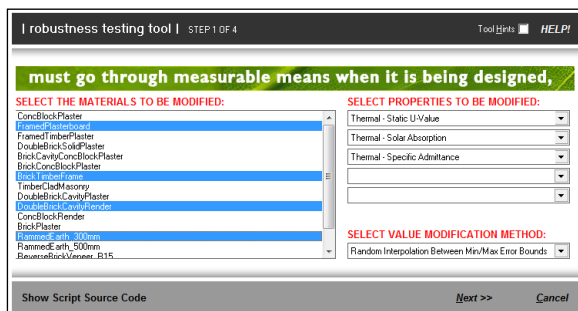

Figure 7 –A custom wizard for comparing results from multiple thermal calculations performed using material properties interpolated from provenance error bands.

Benefits and Disadvantages

As there are no standards or agreed infrastructure for these kinds of processes, Autodesk Ecotect's use of a flexible and customisable database makes it possible to apply the same or similar techniques to zone profiles, operational schedules or other input data on which important calculations depend. Obviously flexibility means custom scripts and minimal interoperability. However, even in its current form, it does provide an environment to develop and test these processes, out of which some standards and agreed infrastructure may eventually grow.

CONCLUSION

This paper has sought to examine how currently available computer tools might work together to build a level of self-documentation of the quality of simulation data that would provide a basis for Quality Assurance in building performance simulation. By creating a specification for the means by which high quality building performance simulation input data was made simply accessible to computer analysis programs like ESP-r and Autodesk® Ecotect™, it becomes feasible to imagine a near future where performance simulation become ubiquitous. Without Quality Assurance systems of this type, that provide systems for making reliable data with high quality provenance available to every designer, community goals that all new buildings should achieve Net Zero Energy status seem mere pipe dreams. Achieving this goal ensures at least that every designer is modelling the physics of the products accurately. Then the focus can turn to the next phase of Quality Assurance: ensuring the design team knows how to use the simulation software appropriately.

REFERENCES

Eisenberg, D., R. Done, L. Ishida. 2002. *Breaking Down the Barriers: Challenges and Solutions to Code Approval of Green Buildings.* Development Center for Appropriate Technology, Tucson, AZ.

Macdonald I A (2002) 'Quantifying the effects of uncertainty in building simulation', PhD Thesis, Glasgow: University of Strathclyde

Yaman H, Tas E, Tancan L (2000) *The content of an ideal web site for building materials information on the world wide web: a Turkish perspective* Digital Library of Construction Informatics: http://itc.scix.net/data/works/att/w78-2000-1069.content.pdf last accessed Jan 2009

Haley, M. (2008) 'Autodesk® Seek' http://www.stress-free.co.nz/autodesk_seek_talk_by_mike_haley last accessed Jan 2009

CBI. http://www.masterspec.co.nz/cbi.asp Last accessed Jan 2009

W3 Corp: http://www.w3.org/TR/wsa-reqs/ last accessed Jan 2009

Google PageRank: http://www.google.com/corporate/tech.html Last Accessed, Jan 2009

Google Web History http://www.google.com/support/accounts/bin/answer.py?hl=en&answer=54068 Last accessed, Jan 2009

[ii] http://windows.lbl.gov/materials/IGDB/default.htm last accessed, Jan 2009
[iv] http://windows.lbl.gov/materials/IGDB/IGDB_Knowledge_Base.htm last accessed Jan, 2009
[viii] http://seek.autodesk.com/ last accessed, Jan 2009